Deloitte.

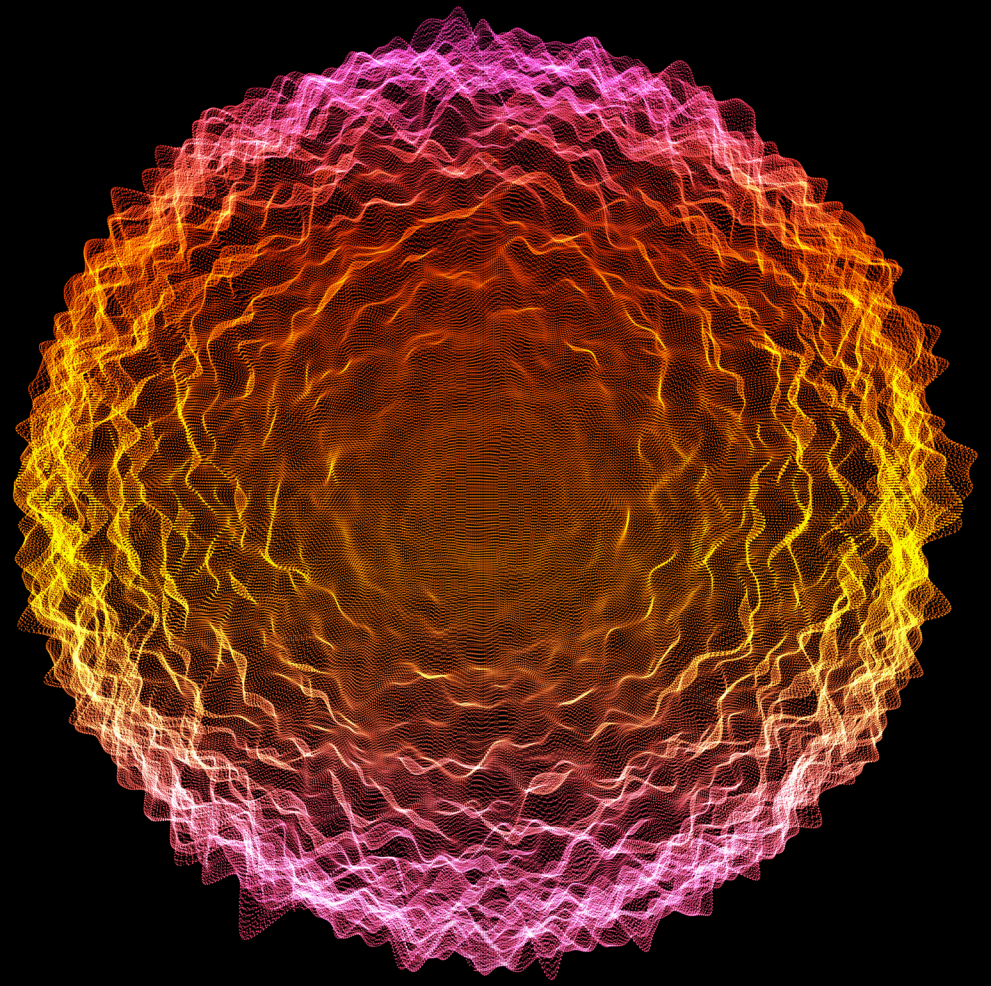
An auditor's mindset in an AI-driven world

An auditor's view on AI governance and risk management within organizations

A point of view by the Deloitte AI Institute

About the Deloitte AI Institute

The Deloitte AI Institute helps organizations connect all the different dimensions of the robust, highly dynamic and rapidly evolving AI ecosystem. The AI Institute leads conversations on applied AI innovation across industries, with cutting-edge insights, to promote human-machine collaboration in the "Age of With". Deloitte AI Institute aims to promote a dialogue and development of artificial intelligence, stimulate innovation, and examine challenges to AI implementation and ways to address them. The AI Institute collaborates with an ecosystem composed of academic research groups, start-ups, entrepreneurs, innovators, mature AI product leaders, and AI visionaries, to explore key areas of artificial intelligence including risks, policies, ethics, future of work and talent, and applied AI use cases. Combined with Deloitte's deep knowledge and experience in artificial intelligence applications, the Institute helps make sense of this complex ecosystem, and as a result, deliver impactful perspectives to help organizations succeed by making informed AI decisions.



No matter what stage of the AI journey you're in; whether you're a board member or a C-Suite leader driving strategy for your organization, or a hands on data scientist, bringing an AI strategy to life, the Deloitte AI institute can help you learn more about how enterprises across the world are leveraging AI for a competitive advantage. Visit us at the Deloitte AI Institute for a full body of our work, subscribe to our podcasts and newsletter, and join us at our meet ups and live events. Let's explore the future of AI together.

www.deloitte.com/us/AIInstitute

An auditor's view on AI governance and risk management within organizations

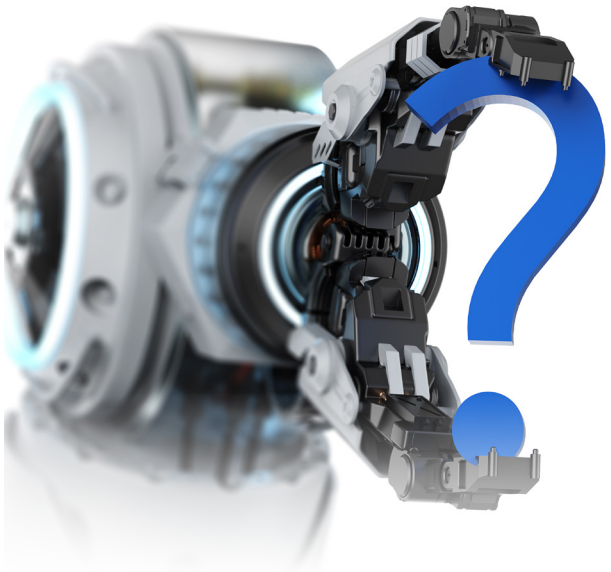
AI is an emerging area that is having a transformational impact on business operations and how organizations achieve their mission, strategy, and objectives.

As this transformation occurs and AI scales within the enterprise, organizations should adapt governance, oversight structures, and processes to promote trust and transparency in AI models. How organizations use the underlying data and resulting outputs can be a game changer, but the reality is that AI models are only as good as the data that feeds them. If the data itself is flawed or changes over time, even subtly, outputs can shift. AI models create predictions, classifications, or new data *based on observable inputs and outputs* as opposed to pre-programmed rules. In other words, AI models *infer* the rules or decision weighting they apply to data. This distinction means that use of AI models can lead to unintended outcomes (e.g., bias, inaccurate decisions or recommendations, etc.).

In response, organizations should adapt their capabilities by testing, interpreting and monitoring both AI models and data to verify that deployed models are operating as intended.

Organizations rely on governance and oversight structures to help ensure effective design and operation of traditional technology applications whether they are simple spreadsheets or sophisticated technology systems. In designing governance and oversight structures, organizations should consider the business purpose, objectives and risks, underlying data, operators, and related applications or processes. While AI applications rely on new and evolving technology, the core concepts surrounding governance, risk, and controls are the same and are still foundational for promoting trust and transparency.

How Auditors Can Help



Auditors are well suited to identify AI-specific risks as well as evaluate relevant governance and oversight structures to help organizations achieve the strategic value they hope to attain using AI.

One of an auditor's core missions is to help enhance trust and transparency through providing assurance on a variety of different subject matters, from financial statements to regulatory compliance. The foundation for auditors to deliver assurance includes evaluating governance, risks, and processes that are relevant to the selected subject matter. An auditor's independent mindset and focus on risk assessment are core underlying concepts for evaluating oversight and effectiveness of AI models.

Like many other stakeholders, auditors are adapting their tried and tested approach to an AI-enabled world. Example considerations include but are not limited to:

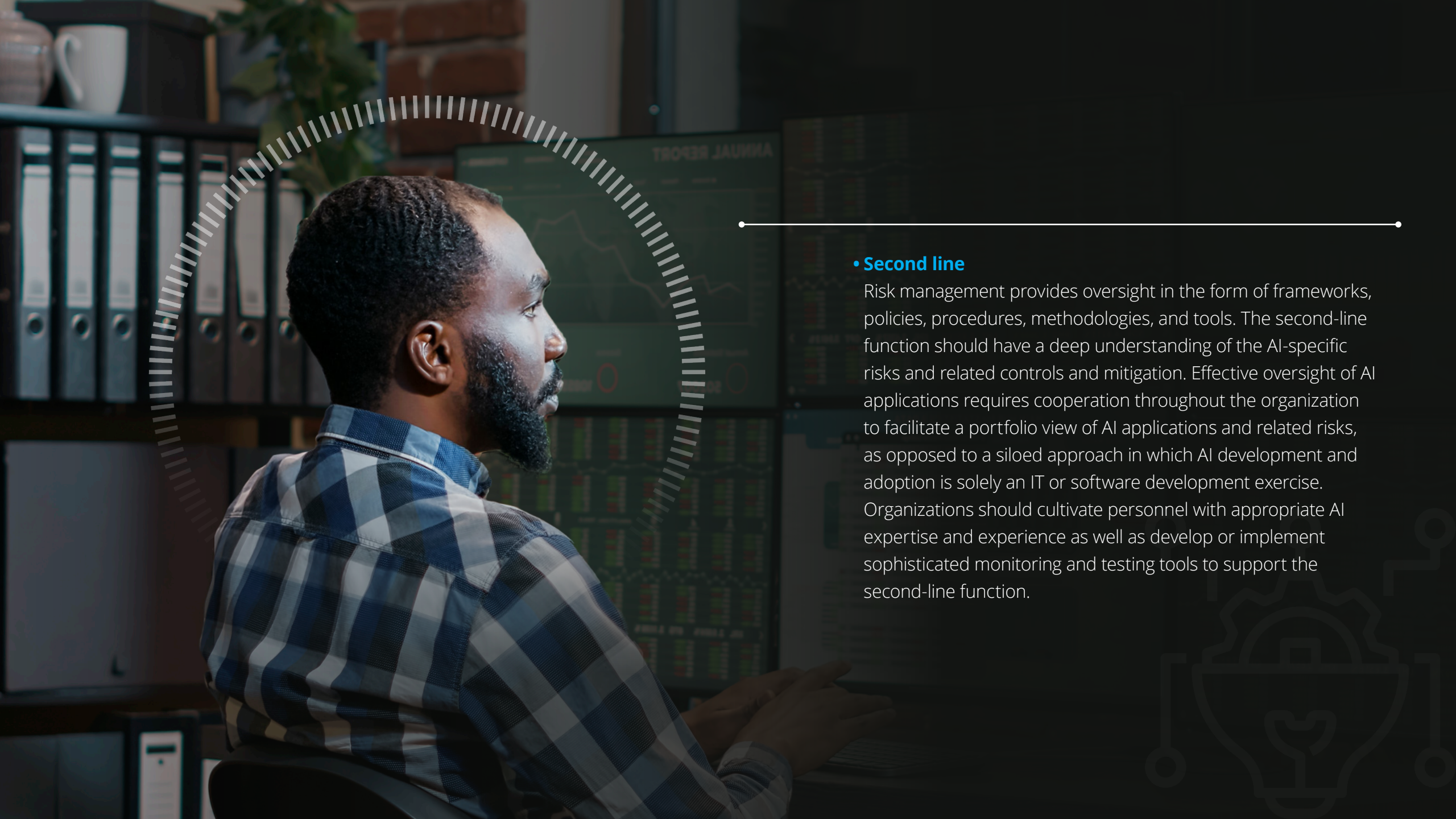
- Are an organization's AI objectives consistent with its mission and strategy?
- How does an organization assess risk and impact of its AI applications as well as consider effects from related core areas (e.g., data management, cybersecurity, etc.)?
- Does an organization's structure and controls lend itself to an effective framework for governance and oversight over its AI model population?
- How does an organization design and implement an effective testing regime for its AI models?
- How does an organization interpret the results of its testing regime and respond to findings or exceptions?

The considerations listed above illustrate the need to think broadly to comprehensively support trust and transparency beyond merely validating the underlying functionality of AI applications. Many organizations across a range of industries have adopted a three line-of-defense model to address these broad considerations. These organizations incorporate various skillsets (e.g., quantitative/technical, governance/policy, internal controls) across the three lines. The evaluation and monitoring of AI risks can be similar to those of other risk types, such as model risk and operational risk. As such, many of the professionals across the existing three lines can by leveraged to effectively oversee, evaluate, and monitor AI applications and underlying models.

An Auditor's View on Three Lines of AI Defense

- **First line**

  Management (process/model owners) has the primary responsibility to own and manage risks associated with development and day-to-day operational activities. Management should have a baseline understanding of risks in AI applications and where they manifest themselves in the specific models and data relevant to the organization's use cases. As part of this understanding, model owners should know how their models work and consider interdependencies between AI applications, how the underlying data is obtained and used, and how the data and the model(s) are secured. It is important that model owners not only test at the development stage by continue to evaluate the performance of their AI applications through ongoing performance monitoring and targeted testing.



- **Second line**

  Risk management provides oversight in the form of frameworks, policies, procedures, methodologies, and tools. The second-line function should have a deep understanding of the AI-specific risks and related controls and mitigation. Effective oversight of AI applications requires cooperation throughout the organization to facilitate a portfolio view of AI applications and related risks, as opposed to a siloed approach in which AI development and adoption is solely an IT or software development exercise. Organizations should cultivate personnel with appropriate AI expertise and experience as well as develop or implement sophisticated monitoring and testing tools to support the second-line function.

- **Third line**

  Internal audit assesses the first-line and second-line functions and reports on its design and operational effectiveness to the board and audit committee. In assessing the first-line functions, internal audit should assess whether AI development and monitoring adheres to the organization's policies, best practices for model development and relevant regulations (e.g., General Data Protection Regulation, Section Five of the Federal Trade Commission Act, etc.). In addition, it should assess that second-line functions are independently reviewing AI risks and corresponding controls. Similar to the controls environment for many other business processes, internal audit should be determining whether AI risks are being properly addressed be the first and second line.

An Auditor's Role

Auditors currently work with each line of defense, senior leadership, and those charged with governance (e.g., board of directors) to assess the organization's control environment. As organizations adopt and expand their use of AI, auditors can have a key role in helping organizations identify and understand AI-specific risks.

In addition to considering the impact of AI applications within the enterprise, organizations should consider external stakeholders, including regulators and investors. As a result of being regulated and helping organizations with compliance, auditors have extensive compliance experience and maintain a close dialogue with a variety of regulators. Additionally, auditors play a critical role in bolstering investor confidence and trust through helping to address investors' expectations for transparency. An auditor's experience with regulators and understanding of regulatory intent and investor expectations, coupled with their independent mindset, can provide valuable insight to organizations as they respond to increased public concern and scrutiny regarding transparency and functionality of AI applications.

Auditors are responding to the proliferation of AI in the business environment. A key aspect of auditors' response is adapting existing and developing new capabilities to assist organizations in promoting trust and transparency in their use of AI, which can support increased and accelerated adoption of AI to help achieve an organization's strategic objectives. **People, process, and technology** are key elements for both auditors and organizations to support trust and transparency in use of AI models:

People

Organizations should leverage existing skill sets in risk management and controls, and model risk management, and then augment those functions with AI specialists (e.g., model owners, data scientists, and developers, etc.).

It's also important to educate employees across the organization and board members. Employees who interact with or are affected by AI applications should have enough knowledge to challenge and check them as well as understand the governance policies and guidelines. Board members should understand situations with elevated AI risks so that they can carefully consider those risks. Auditors factor in a variety of stakeholders in their assessments, from developers and management, to those who are charged with governance and oversight.

Process

Organizations should adapt a governance model to include leading practices, relevant frameworks and regulations that incorporate AI-specific considerations. When considering risks, organizations should have a position that's proactive on preparedness – taking steps to mitigate risks proactively as opposed to reactively.

The adoption of an appropriate framework, such as Trustworthy AI™, is important in addressing AI-specific risk considerations like fairness, bias, and explainability. However, not all AI applications have the same inherent risks. For example, an AI application that helps approve or decline retail bank loans in compliance with lending regulation may have increased risk compared to other applications. Regardless of how AI is used, an effective framework should address both traditional risks and incorporate the unique considerations for AI.

AI-specific considerations based on Deloitte's Trustworthy AI Framework include:

Fair and impartial

Organizations should have mechanisms to enable and test for equitable outcomes across all participants. This includes evaluating AI models for various measures of fairness versus prediction accuracy to identify systematic unwanted bias hidden within AI models.

Robust and reliable

Organizations should strive for AI applications that produce consistent and reliable outcomes. This includes subjecting AI models to stress tests and probing classification boundaries to evaluate AI model stability, resilience, reliability and reproducibility.

Privacy

Organizations should support data privacy. Examples of supporting capabilities include anonymizing data sets and evaluating them for risk of re-identification as well as generating synthetic data sets.

Safe and secure

Organizations should protect AI applications from potential risks (e.g., cyber risks). This includes profiling and detecting exploitable weaknesses (e.g., adversarial attacks through altering data or introducing incorrect data, etc.) in AI models.

Responsible and accountable

Organizations should have structures and policies in place that can help clearly determine who is responsible for the development, operation, and monitoring of AI applications.

Transparent and explainable

Organizations should strive to understand, as best as possible, key drivers that lead to AI outcomes or decisioning. This can include sensitivity analysis to better understand the weights and importance of inputs.

Technology

Organizations should use cutting-edge tools, including AI and data science platforms, to facilitate controlled processes for model development, deployment and ongoing monitoring of performance.

AI and data science platforms can help automate certain aspects of model development by creating standardized workflows for AI models. Examples of these aspects include data extraction and transformation, change management and real-time testing and monitoring of key performance indicators. Technology controls (e.g., change management, access, etc.) are fundamental for any organization, however, robust technology controls are vital for AI applications that typically rely on large, complex and interconnected digital infrastructures. Integrating robust technology controls and enhanced automation for AI is commonly referred to machine learning operations (MLOps). MLOps is a growing practice that aims to standardize and automate many aspects of the AI model lifecycle, thereby reducing errors arising from ad-hoc and manual processes.

MLOps is an important component of the foundation for organizations to integrate supporting tools and functionality to build trust and transparency for their AI applications. Omnia's Trustworthy AI Module showcases how certain tools and functionality can be used to support testing for bias, resilience and reliability, transparency and other aspects of the Deloitte Trustworthy AI framework.

The services described herein are illustrative in nature and are intended to demonstrate our experience and capabilities in these areas; however, due to independence restrictions that may apply to audit clients (including affiliates) of Deloitte & Touche LLP, we may be unable to provide certain services based on individual facts and circumstances.

Authors

Brian Cassidy
US Audit & Assurance
Partner
Deloitte & Touche LLP
Email: bcassidy@deloitte.com

Ryan Hittner
US Audit & Assurance
Managing Director
Deloitte & Touche LLP
Email: rhittner@deloitte.com

Brian Crowley
US Audit & Assurance
Senior Manager
Deloitte & Touche LLP
Email: brcrowley@deloitte.com

Zach Bowman
US Audit & Assurance
Senior Manager
Deloitte & Touche LLP
Email: Jambowman@deloitte.com

John Fogarty
US Audit & Assurance
Senior Manager
Deloitte & Touche LLP
Email: johfogarty@deloitte.com

Deloitte.